

Europäisches Patentamt  
European Patent Office  
Office européen des brevets



(11) EP 0 847 178 A2

(12) EUROPEAN PATENT APPLICATION

(43) Date of publication:  
10.06.1998 Bulletin 1998/24

(51) Int Cl.<sup>6</sup>: H04M 3/56, H04L 12/18

(21) Application number: 97308850.3

(22) Date of filing: 04.11.1997

(84) Designated Contracting States:  
AT BE CH DE DK ES FI FR GB GR IE IT LI LU MC  
NL PT SE  
Designated Extension States:  
AL LT LV MK RO SI

- Harrison, Collin George  
Brookfield, Connecticut, 06804 (US)
- Stern, Edith Helen  
Boca Raton, Florida, 33431 (US)
- Toohey, James J.  
Boca Raton, Florida, 33434 (US)
- Willner, Barry Edward  
Briarcliff Manor, New York 10510 (US)

(30) Priority: 06.12.1996 US 761249

(71) Applicant: INTERNATIONAL BUSINESS  
MACHINES CORPORATION  
Armonk, NY 10504 (US)

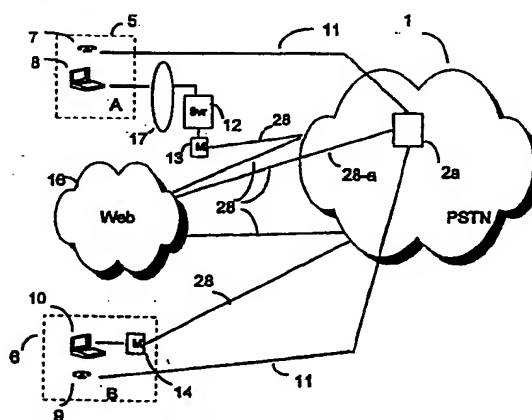
(74) Representative: Zerbi, Guido Maria  
Intellectual Property Department,  
IBM United Kingdom Ltd.,  
Hursley Park  
Winchester, Hampshire SO21 2JN (GB)

(72) Inventors:  
• Dunn, James M.  
Florida, 33435-6215 (US)

(54) Multimedia conferencing using parallel networks

(57) A conference server system, for internal use in the public switched telephone network (PSTN), links to public data communication networks (e.g. the Internet) for distributing computer displayable data between participants in voice telephone conferences. The server system includes a computer sub-system for storing data generated by the participants and distributing the data in coordination with voice presentations of respective participants. Data so stored and distributed includes computer-displayable data. Other data handled by the computer subsystem represents commands and functional requests issued by participants which pertain to the handling of computer-displayable data and voice parameters of the conference. Conference participants, having separate and concurrent access to both the PSTN and the data network, receive and view computer-displayable data prepared by one of the participants in coordination with the respective voice conference. The conference server system connects to voice conference management resources of the PSTN in a manner enabling the server system to control voice parameters of a voice conference in response to data requests received from participants (e.g. requests from hearing-impaired participants for enhanced audio service), as well as enabling the conference management resources to control functions performed in respect to distribution of displayable data in response to spoken requests uttered by the participants, when the telephone system contains appropriate speech recognition equipment.

Figure 3



EP 0 847 178 A2

agement facilities for managing operations in said conferencing circuits to provide conference connections between end users requiring said connections; said conference server system comprising: a computer system for storing and processing data; means connecting said computer system to a data network external to said PSTN for transferring data between said computer system and end users of said PSTN currently participating in a voice conference utilizing said conferencing circuits; said data being transferred in parallel over said data network while signals representing voice utterances of said participating end users are transmitted through said conferencing circuits; said data including displayable image data and data representing functional requests and commands issued by individual said participating end users and requiring attention by said conference server system; said functional requests including requests from a said participating end user to said conference server system to have said conference server system distribute displayable image data, previously furnished to said conference server system by the participating end user issuing the request, to others of said participating end users; and means linking said computer system to said conferencing circuits and said conference management facilities for enabling said computer system to coordinate distribution of said displayable image data between said participating end users in coordination with voice interchanges between said participating end users; said linking means also being useful for enabling said computer system to participate in satisfaction of functional requests and commands issued by said participating end users and said conference management facilities.

During a multimedia conference, voice (or other audio) signals pass between conferees through conventional switch circuits and conference bridge circuits that are internal to the PSTN and that are administered by conventional conference management resources of the PSTN. However, data signals in the same conference (image data and other data) may be routed via a conference server over a public data network external to the PSTN, like the Internet or Web. Each such conference server is administered and controlled by the PSTN and has links to conference connection and conference management resources of the PSTN, such that a conference participant making a voice presentation can furnish image data to the conference server during or prior to the presentation, have the data stored in the conference server, and then have the data uploaded from the conference server to the other participants in precise coordination with parts of the voice presentation referring specifically to respective display images. Connections between conference participants and public data networks linked to a conference server may be dialled (tone keyed) telephone system connections or direct connections through e.g. privately maintained lines and/or networks.

Each conference server and its aforesaid links with-

in the PSTN is also useful to allow conferees to control operational services within both the PSTN and the external data network, both before and during a conference.

In respect to coordinated transfer of image data, a conference server storing data furnished by a conference participant can be requested by that participant to distribute the data to other conference participants. The request can be sent through either the external data network (e.g. as a data signal representing the request) or through the PSTN conference connection as spoken commands if the PSTN is suitably equipped for detecting and reacting to such commands.

In respect to user control of conditions or services internal to the PSTN, such control can be used to interrupt a speaker in the conference; e.g. by a request over the data network causing the conference server to have images displayed to conferees (including the speaker) modified to show an indication requesting interruption of the speaker (e.g. a displayed indication connoting QUESTION TO ASK or NEED TO TALK); thereby allowing for orderly interruption of speakers without having plural participants trying to speak at the same time.

Another use of this facility would be to allow a conferee to request the conference server to adjust amplitude or other characteristics of audio signals currently being received by that conferee; whereby, for instance, a hearing impaired conferee could raise the volume of their received audio without affecting audio volume received by other participants.

Another use would be to allow a conferee to invoke a new "follow me" type of service during a currently active conference wherein e.g. the respective conferee's connection to the conference could be shifted/switched from a fixed line in the PSTN (e.g. at a respective office or residence) to a mobile radio link. Thus, a conferee could leave home or office to travel elsewhere without disconnecting from the conference and without affecting continuity of the conference.

Another use, if the PSTN conference centre is suitably equipped with speech recognition apparatus and applications therefor, would be to allow a voice command over the PSTN conference connection to be detected by the speech recognition apparatus and used to server to control adjustment of a service condition in either the voice or data path or even in both paths (e.g. by displaying an indication to all participants stating or meaning LET'S TAKE A BREAK).

Another aspect of this arrangement is that the conference server can be configured and connected to support inclusion in a conference of "voice-only" and "data-only" participants, in addition to "voice and data" participants (the latter speaking and listening to voice, and sending and receiving data). Thus, a "voice-only" participant (e.g. a telephone user either without a computer or without concurrent access to the data network) could listen and speak to other participants, and a "data-only" participant (e.g. with access to the data network but ei-

URL may be used by "browsing" software, at locations of Web end users, to designate a file to be retrieved from a remote server on the Web.

- Conference call (or session) represents a voice connection service provided by the PSTN wherein two or more end users are connected so that each can speak and the others can hear what is said by the speaker. Participants in a conference call are herein termed participants, attendees, conferring parties or conferees.

## **2. Explanation of Problem Presently Solved**

This invention concerns multimedia conferences wherein conferees concurrently may converse and exchange data, the latter including displayable text and graphic images. In today's facilities for multimedia transmission, data and voice are routed over separate paths such that it is difficult if not impossible to have presentations of image data coordinated with vocal presentations in a coherent manner.

As exemplified in Figure 1, a typical contemporary (prior art) facility of this kind routes voice signals exclusively through the PSTN, and data signals exclusively through a parallel public data network like the Web, but without any possibility for interaction between the two networks; e.g. to allow for coordination of display image presentation with vocal presentation, or to allow for data signal communications to alter services currently effective in the PSTN, etc.

The PSTN, shown at 1, has conference call managing and administering facilities 2 which oversee voice signal connections between conference participants but have no tie-in to simultaneous data transmissions between the participants. In figure 1, A and B represent two (of two or more) end-user stations currently having a conference connection on the PSTN. Equipment used at these stations are indicated respectively at 5 and 6. Station A's equipment 5 includes a telephone 7 and computer/display apparatus 8, and station B's equipment 6 includes a telephone 9 and computer/display apparatus 10. Telephones 7 and 9 are connected through transmission lines 11 in the PSTN and (not shown) switches administered and controlled by the PSTN. Telephones 7 and 9 are not locally connected to respective computers in this illustration, but even if so connected neither the PSTN nor the Web could act to coordinate or otherwise inter-relate voice and data transmissions; i.e. only the end users of these stations could act to coordinate data transfers with voice, by agreeing on a set of protocols beforehand, and such actions could not affect services provided in either network.

To exchange data in the typical arrangement of Figure 1, conferees at stations A and B link to a server 12. The data is transferred e.g. through modems indicated at 13 and 14, transmission line paths shown collectively at 15, and a data communication network such as the

Web shown at 16. In the illustrative example of this figure, server 12 and modem 13 are shown as locally accessible to station A and other not-shown stations via a LAN 17, and remotely accessible to station B via the Web 16 and modem 14. Thus, in the illustrated example, station A and the not-shown other stations could be computer stations within a building or a group of geographically proximate buildings (e.g. a building or buildings locally serving a business enterprise), and server 12 and modem 13 could be shared via LAN 17 by occupants of the respective building(s). Thus, the occupants would receive shared data handling and modem conversion services, rather than requiring separate data handling and modems at their individual offices or sites. In contrast, modem 14 is used only by station B and is within a few feet of the computer at that station.

In this configuration, modems such as 13 and 14 are required only for linkage to the applicable data network such as the Web. Also, server 12 can be located anywhere within communication reach of the respective set of stations that it serves (i.e. a WAN could be substituted for LAN 17 if necessary and/or appropriate). Furthermore, it should be understood that server 12 may be only one of plural servers through which the conferees exchange data. Although lines 15 between modems 13 and 14 are shown as passing through both the Web 16 and PSTN 1, in this example they have no connectivity to the call management facilities 2 of the PSTN.

Referring to Figure 2, the prior art call management facilities 2 include a call management system 20 (e.g. a computer system or the like) which manages and administers conventional switching apparatus 21 and conventional bridging apparatus 22 to establish and terminate conference call connections. Apparatus 22 comprises facilities allowing for voice signals to be amplified, summed and distributed among the conferees.

In this environment, functions and end user services provided by the Web are not affectable by voice signals in the PSTN, and conditions in the voice connection path are not affectable by signals sent through the data connection/Web.

Figure 5 shows how a conference session is established and maintained in the prior art environment shown in Figures 1 and 2. As shown at 23, each participant establishes separate linkages to the PSTN and to a server such as 12 in the data network/Web, the former exclusively for voice contact and the latter exclusively for data interchange. Therefore, each participant requires simultaneous access to the PSTN and the data network via at least two physically or logically separate lines. Typically, in this environment, arrangements for the conference are made in advance and may require assistance of telephone system operators to link individual callers to the conference and announce their presence to those already participating. Usually, the advance arrangements define a maximum number of attendees, and once that number is reached no additional participants are permitted (unless and until one or more

and/or invoked by conferees during a conference.

### 3. The Present Solution

An overview of how the present invention solves the foregoing problems is presented in figures 3, 4, 6 and 6a.

Figure 3 shows the arrangement of Figure 1 modified in accordance with our invention. Elements functionally identical to those in Figure 1 are denoted with identical numbers. Consequently this shows that novel elements of the invention are in the conference control centre 2a and are associated at least in part with the data path represented by lines 28, 28-a.

Figure 4 shows that conference centre 2a includes a conference management system 20a and a conference server system 30. Conference server 30 is a communications adapted computer system maintained and administered by the PSTN. It's loose integration with call management resources of the PSTN, including its connection 31 to conference management facilities such as 20a, enable it to control services performed relative to an existing PSTN conference connection, in response to signals sent over a path external to the PSTN; exemplified by the indicated external connection(s) 32 to the Web (these connections being viewable as extending to line 28-a in Figure 3).

These external connections permit new conferees to easily be added to an active conference without assistance of telephone company operators. Such added conferees may be pre-scheduled to join the conference at its inception, or they may join on an ad hoc basis, and their additions may be displayed to prior participants eliminating need for announcements by assisting telephone operators. They also enable conferees to easily and economically control other PSTN conference-useful services (either before or during a conference) without operator assistance; e.g. to vary connections of participants moving between fixed desk and mobile cellular telephones (in a manner further discussed later on).

Conference management system 20a differs from "prior art" conference management system 20 of Figure 2 primarily in its connection 31 to conference server 30, and in functions and services derivative from that connection. As discussed later, speech recognition apparatus and associated applications within the PSTN may be linked to conference bridge 22, for recognizing voice commands of conference participants and enabling either management system 20a or the data network, or both, to perform actions associated with such commands; e.g. real time actions instantly affecting conditions in either the voice connection path, the data transmission path, or both.

It will also be shown later that although participating stations in the arrangements of Figures 1 and 3 have separate telephones and computers, as well as separate communication paths for each, the conference centre arrangement of Figure 4 would easily be adapted to

serve conferees having only a telephone, or only a computer, or conferees having both a telephone and computer operating through a single physical line connecting to either the PSTN or the Web. Adjustments needed to support these reduced configurations are discussed later in this description. Furthermore, although station equipment external to the PSTN and Web is shown in Figure 3 as identical to that in Figure 1, it should be understood that this is done only to facilitate comparison of the present inventive solution to prior art. Therefore, it should be understood, and will become apparent from later parts of this description, that stations having many different forms of equipment are supportable by the present arrangement.

Figure 6 provides an overview of capabilities afforded by the present conference server arrangement.

As indicated at 36, conference participants in both voice and data aspects of a conference still establish parallel voice and data connections over physically or logically separate lines extending to the PSTN and data network. However, as indicated at 37, data transferred between the conferees is funnelled through conference server 30. Consequently, servers external to the PSTN may not be required for presently contemplated conference services, except perhaps for providing data collection, storage and distribution services within e.g. private enterprises (office buildings, plants, etc.). In other words, end users at self-contained participating stations like station B, having unshared data handling and modem facilities, need only link to PSTN conference servers such as 30 and need not have any direct links to other servers such as 12.

As suggested at 37, since server 30 has direct links to conference management facilities of the PSTN, image data transmissions on the Web are coordinatable with vocal aspects of the conference, and other data transmissions may be used by participants to alter or invoke services affecting their voice connections (e.g. volume of voice signals sent to them, interruption of speakers, switching phone connections e.g. from fixed lines to mobile radio links, etc.). Block 37 also indicates that data security may be provided by the conference server at "relatively small cost"; since that server and its functions are effectively shared by all users of the respective PSTN conference management and bridging facilities, rather than by potentially fewer users of (possibly a larger number of) servers such as 12, whence its security functions may be apportioned among a potentially larger population of users than that of servers such as 12 external to the PSTN. This should be more fully understood from the description below of Figure 11.

Figure 6a elaborates on functions indicated in Figure 6. Voice signals flow only through switch and conference bridge connections in the PSTN (block 40). Data signals flow only through data connection paths between the conference server and conferees, generally through the Web, and generally over connection paths separate from the voice connection (block 41). Voice

the image currently displayed to the participants; e.g. an icon representing ASK QUESTION, LET ME TALK, etc. This may be done either by signalling browsers at participant stations to modify the portion of the displayed image containing the icon, or by signalling that the current image is "stale" and requires reloading (in which case the browsers would request reloads and receive the current image with the icon illuminated), or by any other action resulting in an appropriate visual cue to the speaker that another party wants to speak.

Figure 10 illustrates another example of action pertaining to the voice conference. Here, the request (operation 64) is to alter a characteristic (e.g. volume) of voice signals that are being sent to the requestor (e.g. a hearing impaired party). In this case, the conference server forwards a corresponding request to the conference manager (operation 65), and the latter takes action 66 relative to the voice path (conference bridge, etc.) to effect the desired change. A similar request action could be used to have the conference server and conference manager change the phone line connection of a requesting participant; e.g. to switch that participant's link to the conference from a fixed line (residential or office) to a mobile radio link. In this last example, the switching action could be implemented by having the conference manager first communicate with a mobile switching office to establish the mobile radio link to the conference bridge, and then act to sever the connection of the requestor's fixed line.

Figure 11 schematically illustrates a conference control system, in accordance with the present invention, which shows the different configurations of user stations that are supportable and how such support is given. As such, this may be considered as showing the best mode contemplated for carrying out our invention.

This figure shows four user stations -- stations A, B, C and D -- having various different configurations of equipment which for the following discussion are assumed to be actively connected, via conferencing centre 70 in the PSTN, in a multimedia conferencing session (i.e. users of these stations are participants in that session).

Station A contains only a single analog telephone 72; i.e. its user is a voice-only participant in the conference. Station C contains an analog telephone 74, digital computer 76 and modem 78, with the telephone connecting to conference centre 70 in a first connection path within the PSTN that is described below, and the computer and modem connecting in a physically separate second path within the PSTN, also described below; i.e. its user is a participant in both voice and data elements of the conference, whose voice and data flow through physically different paths within the PSTN. Station D contains only a computer 80, modem 81 and speaker 82 and microphone 84 both attached to the computer; i.e. its user is a voice and data participant whose voice and data are constrained to pass through the respective computer 80 and modem 81, and there-

fore have only a single physical path of connection to the conference centre (which path, in the illustrated arrangement, is entirely within the PSTN). Finally, station B contains an analog telephone 85 and digital computer 88 which connect to the conference centre through physically separate paths described below, one in the PSTN (the voice path) and the other external to the PSTN (the data path).

Computers at these stations (those other than station A) are shown as laptop style personal computers, but it is understood that they may be computers of any physical form (e.g. laptop, desktop, floor console, or even computers integral to an appliance such as a TV monitor), and need not even be personal type computers (i.e. they could even be computers like mainframes or such that are larger in size and operating capabilities than personal computers). The only requirement is that they contain browser software that is compatible with applets or scripts written in a common language such as Java, and that communicate with protocols that are compatible with those employed at the conference server to be described below.

Conference centre 70 comprises call connecting circuits 90 containing switching circuits needed for simple non-conference connections as well as bridging circuits required for conference connections. The size and capabilities of this element are dependent upon the anticipated peak traffic in non-conference and conference calls for the respective conference centre.

Conference centre 70 also comprises a call management facility 92, a conference server facility 93, an administrative facility 94, an (optional) speech recognition facility 95 (for speech recognition applications including those relevant to conference management), and an (optional) archive facility 96 (for voice and data archiving applications that could e.g. allow voice and data portions of a conference to be replayed to conference participants at some later date). Facilities 92, 93, 95 and 96 have direct connections to connecting circuits 90 for handling signals to and from end user stations. All of the facilities 92-96 connect to a LAN 100 for communicating with each other. The connection circuits 90, facilities 92-96 and LAN 100 are all components of conference centre 70, and the latter is a component of the PSTN.

Telephone 72 in station A connects to connecting circuits 90 through a part of the PSTN indicated at 104. Telephone 74 in station C connects to circuits 90 through a part of the PSTN shown at 106, while computer 76 and modem 78 in the same station connect to conference server 93 through another part of the PSTN 108 and a modem bank 109 (i.e. this station has two physically separate lines locally connecting to the PSTN and having voice routed through the switch circuits 90 and data routed to the conference server over a path parallel to and separate from circuits 90). The computer 80 and modem 81 in station D connect to modem bank 109 via a link to PSTN portion 108 that is physically separate from that of modem 78 in station C. Consequently, the

- selectively block portions of a conference (voice and/or data) from presentation to some participants; e.g. to present confidential data to participants authorized to receive such and block it from participants not authorized to receive it.

The foregoing and other aspects of the invention will be further appreciated from the following claims.

#### Claims

1. A conference server system, for use in a public switched telephone network (PSTN) having conferencing circuits for routing analog signals, including voice signals, between end users of said network, and conference management facilities for managing operations in said conferencing circuits to provide conference connections between end users requiring said connections; said conference server system comprising:

a computer system for storing and processing data;

means connecting said computer system to a data network external to said PSTN for transferring data between said computer system and end users of said PSTN currently participating in a voice conference utilizing said conferencing circuits; said data being transferred in parallel over said data network while signals representing voice utterances of said participating end users are transmitted through said conferencing circuits; said data including displayable image data and data representing functional requests and commands issued by individual said participating end users and requiring attention by said conference server system; said functional requests including requests from a said participating end user to said conference server system to have said conference server system distribute displayable image data, previously furnished to said conference server system by the participating end user issuing the request, to others of said participating end users; and

means linking said computer system to said conferencing circuits and said conference management facilities for enabling said computer system to coordinate distribution of said displayable image data between said participating end users in coordination with voice interchanges between said participating end users; said linking means also being useful for enabling said computer system to participate in satisfaction of functional requests and commands issued by said participating end users and said conference management facilities.

2. A conference server system in accordance with claim 1 wherein said means for connecting said computer system to said conferencing circuits and conference management facilities includes:

means responsive to certain said functional requests for signalling said conference management facilities to vary conditions of operation in said conferencing circuits.

3. A conference server system in accordance with claim 1 or 2 wherein said PSTN includes: speech-recognition means linked to said conferencing circuits for detecting spoken commands and functional requests uttered by said participating end users; and

means linking said speech-recognition means to said conference server computer system for enabling said conference server computer system to participate in satisfaction of said spoken commands and functional requests; and wherein said conference server computer system comprises:

means responsive to signals received from said speech-recognition means for distributing individual pages of said displayable image data to said participating end users and for performing other functions relevant to said voice conference in which said end users are currently participating.

4. A conference server system in accordance with claim 2 wherein said signals to vary conditions in said conferencing circuits are useful to cause the telephone connection of a said participating end user to be changed so as to enable the respective end user to change between stationary and mobile phones separately connected to said PSTN without interrupting continuity of a currently active conference.

5. A conference server system in accordance with claim 2 wherein said signals to vary conditions in said conferencing circuits are useful to effectively vary the volume of sound currently being delivered to a telephone receiver used by one of said participating end users.

6. A conference server system in accordance with any preceding claim wherein said connections between said computer system and said conferencing circuits are useful to enable said computer system to transfer signals representing voice and data originated by said participating end users between said participating end users, and to enable said computer system to route said voice signals to said conferencing circuits and said data signals to said data network; whereby participating end users lacking

- display additions of new participants and removal of old participants as they occur;
- permit said conference to be sub-divided into plural sub-conferences between subgroups of the participants in said conference; 5
- block portions of a conference relative to a subset of its current participants (e.g. to prevent unauthorized participants for listening to and/or viewing confidential information). 10

15

20

25

30

35

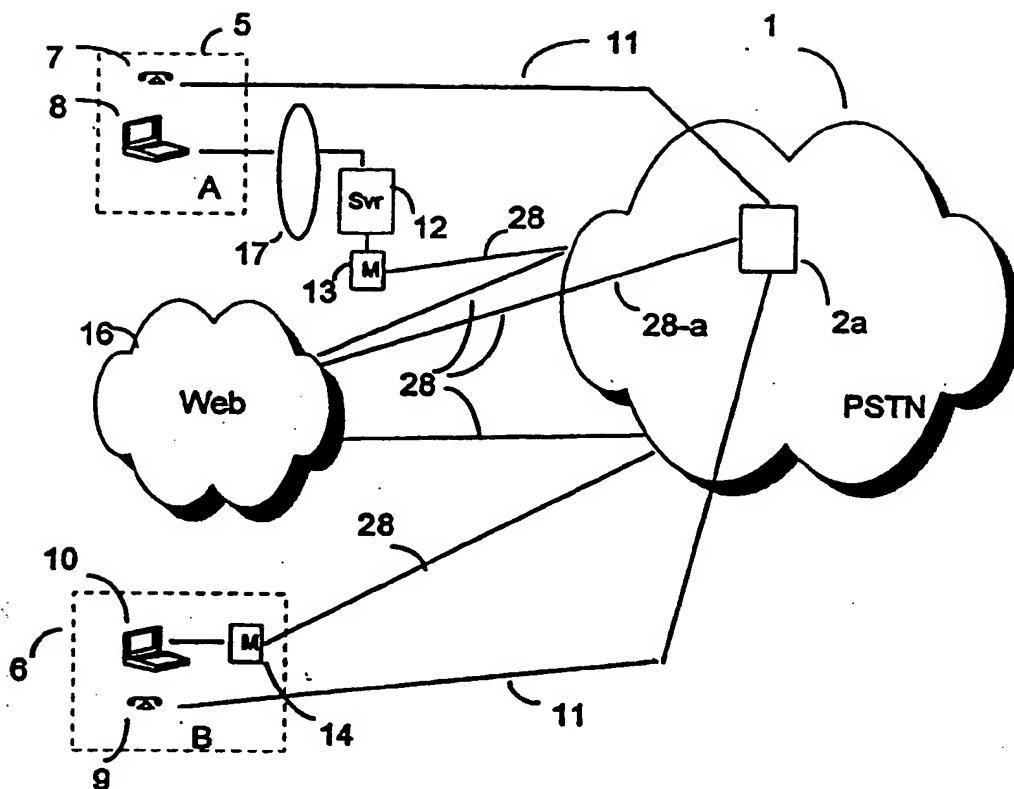
40

45

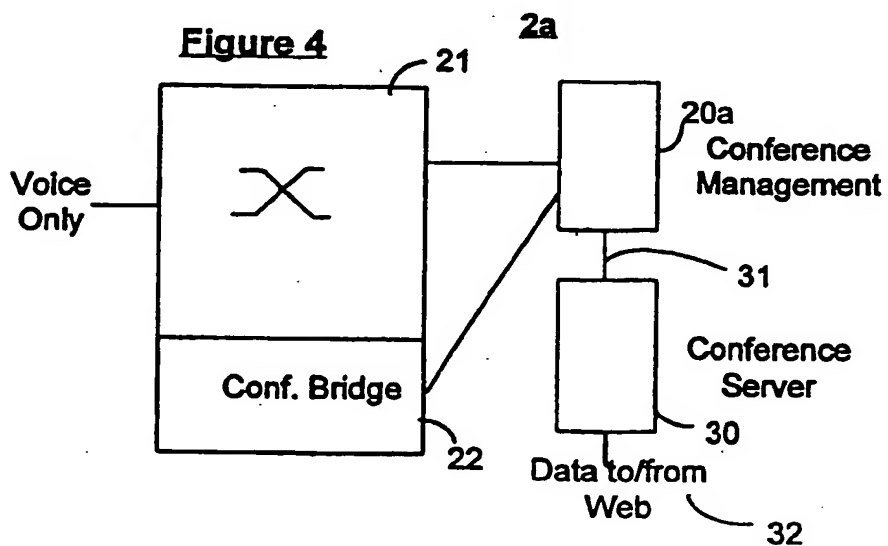
50

55

**Figure 3**



**Figure 4**





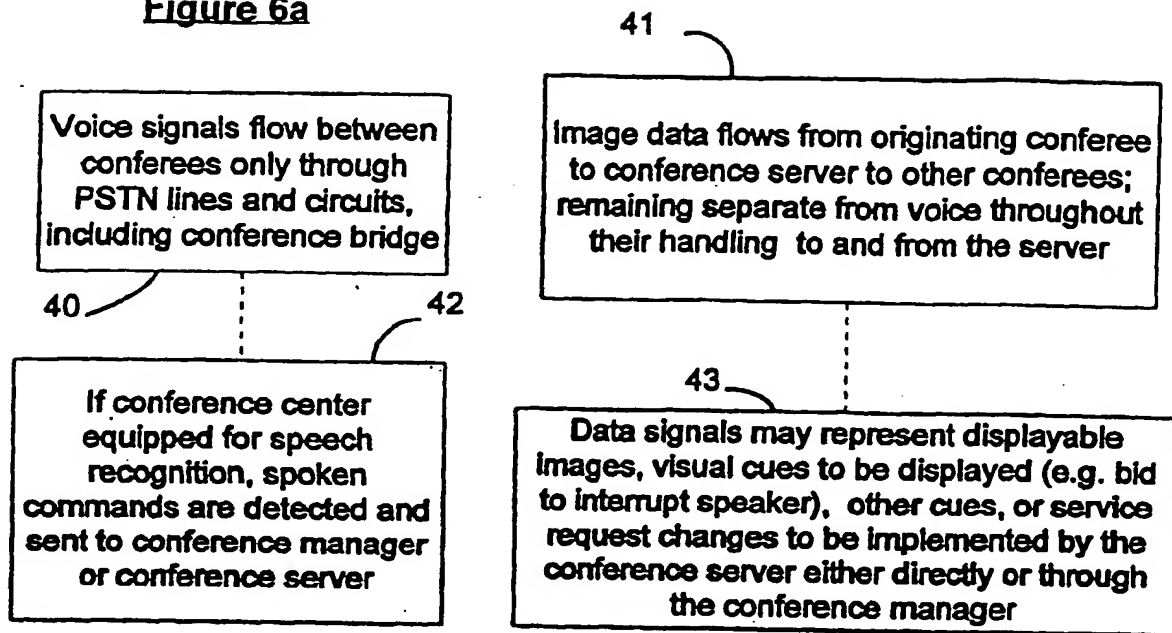
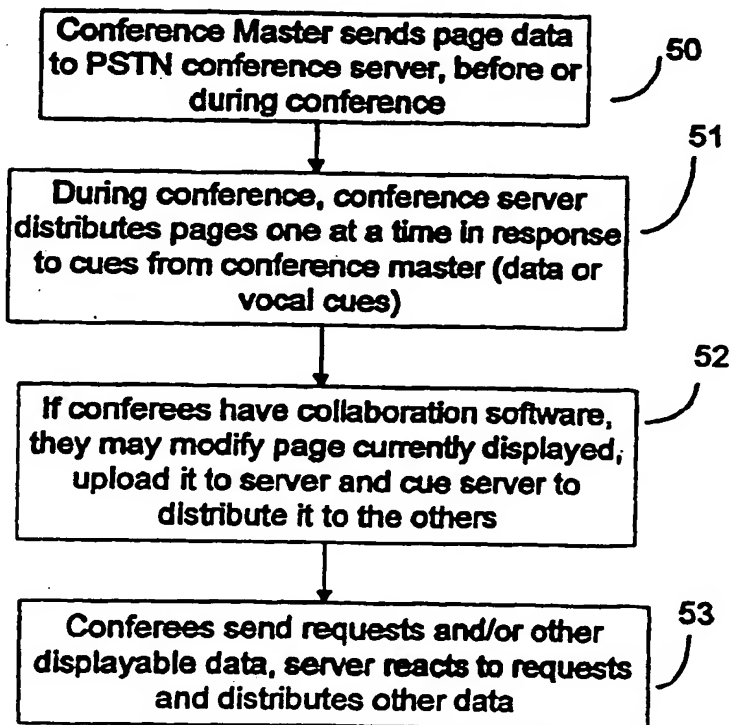
**Figure 6a****Figure 7**

Figure 11

